SPATIAL IMAGE CONTENT BANDWIDTH REQUIREMENTS FOR SYNTHETIC VISION DISPLAYS

Shaowei Yang, Thomas Schnell, and Katherine Lemos

University of Iowa, Operator Performance Laboratory, Iowa City, IA

Abstract

Synthetic Vision Systems (SVS) provide pilots with an unprecedented level of terrain awareness. Questions have been raised as to what level of resolution is needed for the raster based screen on which the terrain is shown. We conducted a flight simulator experiment where we systematically varied the pixel resolution of an SVS from 80 pixels per inch (PPI) to 120 PPI and determined pilot performance in terms of cross track error (XTE) when flying with sole reference to terrain. We then conducted a Fast Fourier Transform (FFT) analysis of the SVS images under the various levels of pixel resolution to determine how spatial image content could be correlated to pilot performance (XTE). We found that the log image amplitude correlated well with XTE and we propose a method wherein we may be able to predict pilot performance based on terrain image content. The method shown in this paper will allow display designers to harmonize the visual spatial bandwidth requirements of pilots with the bandwidth provided by the SVS images.

Introduction

With promising benefits of SVS for flight safety especially under conditions of limited outside visibility, the optimum design features of the SVS display system has been under intensive investigation [1-3]. Much of the focus is on pilot performance and workload as a function of SVS display configuration [3]. Display resolution is one of the potentially influential design features. Displays with higher resolutions are capable of presenting finer image details, and therefore more visual information, presumably resulting in better terrain awareness. The human visual system is limited in perceiving very high spatial frequency, and a threshold level of spatial frequency content might exist for the SVS display beyond which a marginal rate of return may exist for increasing display resolution when it comes to pilot performance. Consequently, a certain optimum bandwidth requirement of spatial image content might exist for SVS displayed images.

The objective of this research was to discover if a link between spatial image content and pilot performance could be found when flying with sole reference to SVS terrain images (no pathway guidance).

Literature Review

Human Visual System

Visual stimuli can be decomposed into amplitude and frequency components using Fourier Transform. The human visual system does not deal with all available spatial frequencies in one single pathway but rather has several pathways (channels), each one of which is optimized to deal with a specific narrow bandwidth of spatial content [4-7]. The human visual system decomposes a visual

stimulus into its spatial consituents in terms of luminous power of spatial frequency channels and perceives components of adjacent spatial frequency channels quasi-independently. The human visual system is limited in perceiving very high spatial frequency. The highest spatial frequency that the human visual system can resolve is around 60 cycles per degree (cpd) of visual angle, and the most sensitive spatial frequency to the human visual system is around 4 cpd.

MTFA Based Image Quality Metrics And Human Performance

Much research has been conducted to measure electronically displayed image quality and to correlate image quality metrics with human visual performance [8]. To date, well-known and successful image quality metrics include: the Modulation Transfer Function Area (MTFA), the Gray Shade Frequency Product (GSFP), and the Integrated Contrast Sensitivity (ICS), to name a few. These three metrics are based on the concept of modulation transfer function (MTF). For a detailed discussion of image quality metrics, refer to [8].

Sinai et al. conducted a series of psychophysical experiments to investigate the effects of images created by different artificial rendering techniques on different visual perceptual tasks [9]. Presumably, image quality depends more on high frequency components. The higher the frequency of an image and the greater the amplitude of such frequencies, the higher the image quality is. Different visual tasks rely on varying visual cues and spatial frequency image components [9]. By using Fourier analysis, Sinai et al. split the spatial frequency spectrum into three channels that were 2 octaves wide: 0.25-1 cpd, 1-4 cpd, and 4-16 cpd, or low, medium and high frequencies, respectively. Their study found that human visual performance is positively correlated to the gross power of the displayed images. We are able to confirm this finding with regard to piloting aircraft with reference to terrain images on an SVS.

Method

Digital images can be transformed from the spatial domain to the spatial frequency domain through a process called Fourier Transform, which quantifies and represents image content mathematically in the form of a spatial frequency and amplitude pair [10]. In other words, each pair of spatial frequencies and amplitudes in the spatial frequency domain corresponds to a portion of the original image of a certain spatial frequency.

The SVS displayed terrain images were subjected to the Fourier Transform for the purpose of finding an image property that characterizes image content and correlates with pilot flight performance.

Experimental Design

In order to being able to correlate image content to pilot performance, it was necessary to conduct a simulator study where we varied display resolution (PPI) and measured cross track error (XTE) as a function thereof. We used a computer generated D-Optimal fractional factorial design in 16 runs as shown in Table 1 [11]. A total of 34 instrument-rated pilots were recruited as subjects to fly the 16 runs in an experiment session of about 3 hours.

Table 1. Experiment Design Matrix

Run Number	Display Resolution (ppi)	Field of View (degrees)	Terrain Type	Terrain Texture	DEM (arc seconds)
1	120	60	Hilly	EC	30
2	90	60	Flat	PR	3
3	105	90	Hilly	PR	3
4	90	22	Hilly	EC	3
5	80	90	Flat	EC	30
6	80	30	Hilly	PR	30
7	90	90	Flat	PR	30
8	105	22	Flat	EC	30
9	105	30	Flat	EC	3
10	105	60	Hilly	PR	30
11	120	22	Flat	PR	30
12	120	30	Flat	PR	3
13	120	90	Hilly	EC	3
14	90	30	Hilly	EC	30
15	80	60	Flat	EC	3
16	80	22	Hilly	PR	3

Note: EC stands for Elevation Color, and PR stands for Photo Realistic.

Two-dimensional Image FFT Analysis

Two-dimensional Fast Fourier Transform (FFT) was used to transform SVS displayed images into the spatial frequency domain. FFT is based on Discrete Fourier Transform (DFT), but is much more efficient than DFT and generates identical results. The DFT of a two-dimensional digital image *f(x,y)* of size $M \times N$ is given in Equation (1)

$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(ux/M + vy/N)}, \quad (1)$$

for x=0, 1, 2,…, M-1 and y=0, 1, 2,…, N-1 [12]. Here x and y are the horizontal and vertical pixel positional indices of an image originating from the upper left corner. The resulting Fourier Transform has the same dimensions as the original image, but in a different domain, the spatial frequency domain. Here, u and v are horizontal and vertical spatial frequency indices, respectively. The actual spatial frequency in pixel per inch (ppi) at u and v can be calculated as

$$f_u = u \Delta u \quad (2)$$

and

$$f_v = v \Delta v, \quad (3)$$

where

$$\Delta u = \frac{1}{M \Delta x}, \quad (4)$$

$$\Delta v = \frac{1}{N \Delta y}, \quad (5)$$

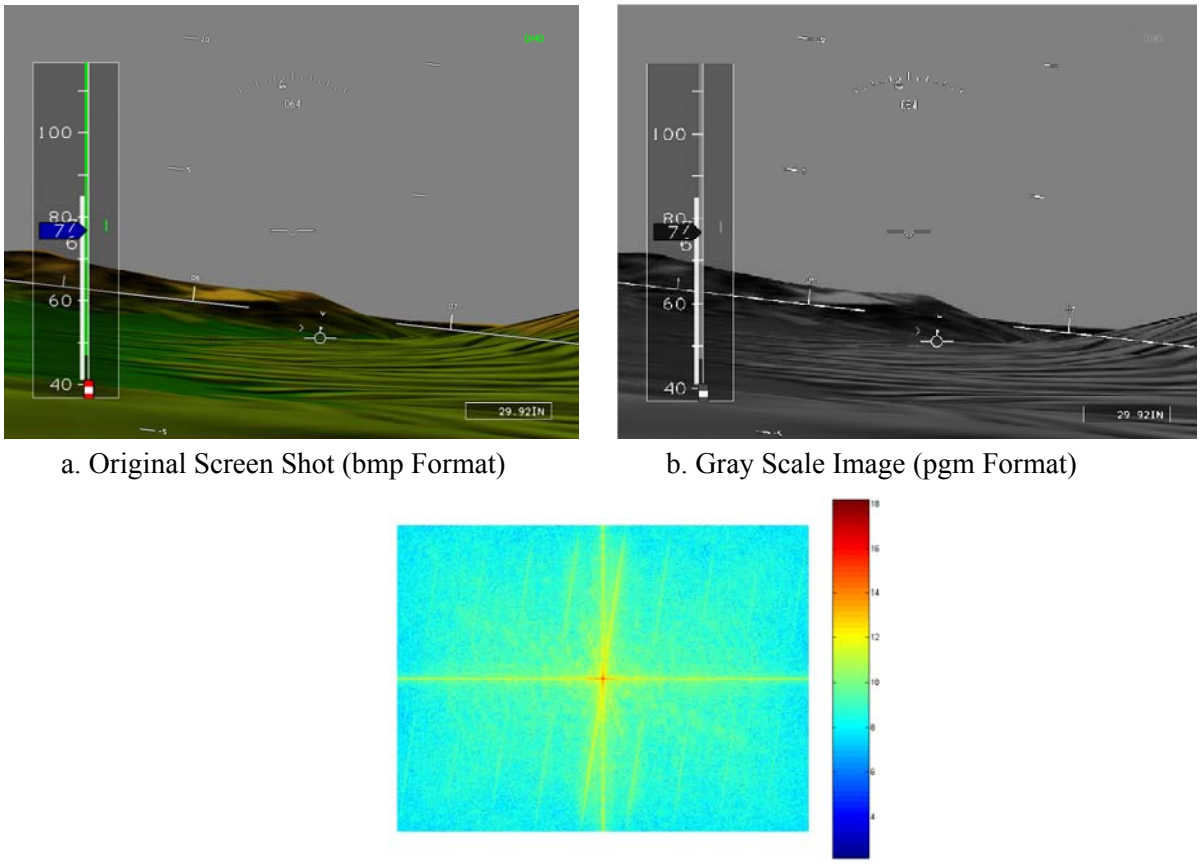
Δx and Δy are the horizontal and vertical image spatial pixel intervals in inch per pixel (ipp).

Screen Shot Generation and Low Pass Spatial Frequency Filtering

A series of screen shots for each run in the experiment (see Table 1) were generated from an autopilot run at GAWSTOO, a fixed base flight simulator (see Figure 1) at the Operator Performance Laboratory (OPL), at The University of Iowa.

Figure 1. GAWSTOO Flight Simulator at OPL

The screen shots were in color bitmap format with a pixel resolution of 1024×768. The numerical representation of a color bitmap is a three-dimensional matrix (two positional dimensions and one color dimension integrating RGB color information). To achieve the pure effect of spatial image content, the color effect was removed by converting the color bitmap images into gray scale images. Portable Gray Map (pgm) format was used since it is a standard bitmap based format providing a maximum of 256 gray scale levels. FFT was implemented on the SVS displayed screen shots in the format of pgm. Figure 2 shows an original screen shot, the gray scale version in the format of pgm, and the graphical representation of the FFT result, in which the spatial frequency components have been shifted so that zero frequency is located at the center and the frequency increases gradually from the center to the outer edges. The pixel resolution of the original screen shots was 1024×768. And the FFT result was a complex numbered matrix with the same dimensions of 1024×768. The center point is called the Direct Current (DC) component, representing 0 frequency and the general brightness level of the image.



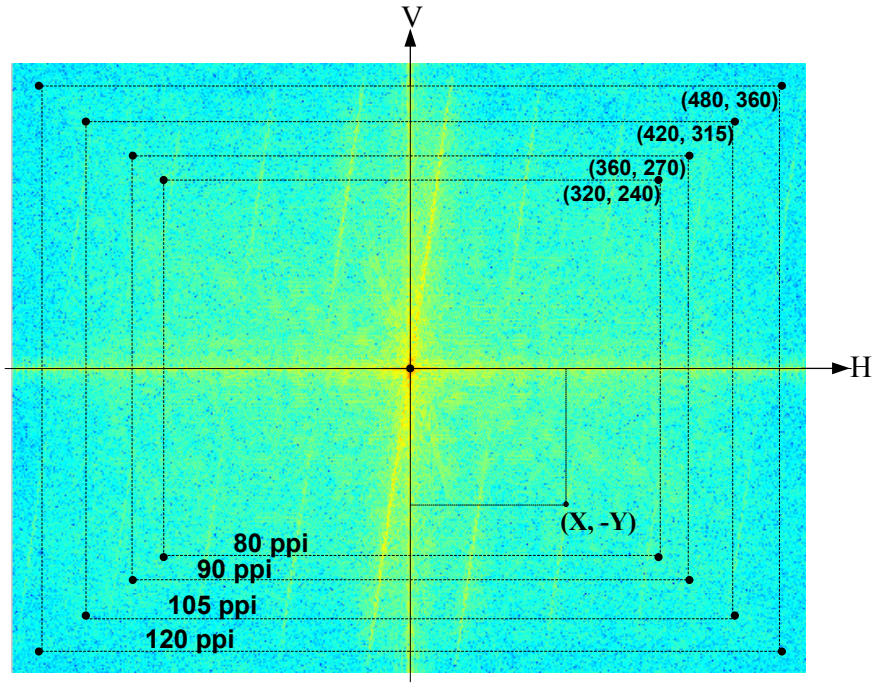
a. Original Screen Shot (bmp Format)

b. Gray Scale Image (pgm Format)

c. FFT Result ($\log_{10}$ Transformed)

Note: The color-coding indicates the $\log_{10}$-transformed amplitude of each point in the spatial frequency domain.

Figure 2. An Original SVS Screen Shot, the Converted Gray Scale Image, and the FFT Result



Note: The axis labeled "H" denotes the horizontal spatial frequency index, and the axis labeled "V" denotes the vertical spatial frequency index. The four frames labeled "80 ppi", "90 ppi", "105 ppi" and "120 ppi" illustrate the structure of low pass filters.

Figure 3. An Image FFT Result in a Cartesian Coordinate System and an Illustration of Low Pass Filters

Since the resolution of the original screen shot (128 ppi) was higher than 80, 90, 105, and 120 ppi when displayed in an 8×6'' area designed in the experiment, the spatial image content of high frequencies needs to be removed to match the actual experimented SVS display resolutions. In Figure 3, a Cartesian coordinate system was superimposed onto the FFT result of the SVS image, as shown in Figure 2. The FFT result was separated into several areas by frames labeled "80 ppi", "90 ppi","105 ppi", and "120 ppi". The area inside each frame represents the image content of a screen shot that could be exhibited using an 8×6'' display with a maximum physical resolution, as labeled. For example, the area inside the frame labeled "80ppi" is the image content that can be exhibited by a display with a maximum physical resolution of 80 ppi and a dimension of 8×6''. Therefore, four low pass filters were used to simulate SVS displayed images of different levels of resolution from the original screen shots.

Spatial Frequency and Amplitude Computation

Cpd is a commonly used spatial frequency unit, usually measured using sinusoidal wave gratings [4]. A pair of bright and dark bars in a sinusoidal wave grating is called one cycle. Cpd is calculated as the number of cycles in a grating object divided by the visual angle subtended by the grating object, as shown in equation (6).

$$cpd = \frac{N}{2\arctan\frac{W}{D}}, \tag{6}$$

where the numerator, N, is the number of cycles in a sinusoidal wave grating, and the denominator is the visual angle subtended by the grating. W is the width of the grating, and D is the distance from the grating center to the middle of the observer's eyes using the same unit as W. We use cpd rather than PPI because the human visual system really does not care so much about the number of pixels in a unit distance but rather how many pixels are subtended in a visual angle. In the cpd notion of spatial resolution, we incorporate the viewing distance from the pilot to the display.

Each point in the spatial frequency domain corresponds to a pair of horizontal and vertical frequencies. Note that the SVS display was kept at 8×6'' throughout the experiment regardless of the display resolution. In this case, Δx is 1/128 (6/768) ipp, and Δy is 1/128 (8/1024) ipp. Therefore, Δu and Δv are 1/6 and 1/8 ppi respectively (see equations (4) and (5)).

One computation example of the spatial frequency in cpd and the amplitude of the FFT result is given below. A point (shown in Figure 3) with an absolute horizontal pixel distance X and an absolute vertical pixel distance Y from the center point in a shifted FFT result has a horizontal frequency in cycles per inch (cpi)

$$f_u = X \cdot \Delta u \tag{7}$$

and a vertical frequency in cpi

$$f_v = Y \cdot \Delta v \,. \tag{8}$$

Note that one cpi is equal to two ppi's. Frequencies were converted from cpi into cpd with the following procedure. The distance (d) from the middle point between the pilot's eyes to the center of the primary flight (PFD) display was measured. Then, the corresponding cpd value of a cpi measure was calculated according to the following equation

$$cpd = \frac{cpi}{2 \cdot \arctan\left(\frac{0.5}{d}\right)} \,. \tag{9}$$

For all experiment runs, the average d was approximately 25 inches. Therefore, the point with a horizontal distance of X and a vertical distance of Y from the center in Figure 3 has a horizontal frequency in cpd

$$f_u' = \frac{f_u}{2.29} \tag{10}$$

and a vertical frequency in cpd

$$f_v' = \frac{f_v}{2.29} \,. \tag{11}$$

The spatial frequency of this point is

$$f = \sqrt{f_u'^2 + f_v'^2} \,. \tag{12}$$

Note that the numerical value at each point in the FFT result is a complex number ($a+jb$). The amplitude is defined as $\sqrt{a^2 + b^2}$.

The amplitude of the DC component at each level of the experimented display resolutions was scaled according to pixel resolution dimension ratios relative to the original screen shots in order to maintain the original brightness for all experiment scenarios.

Total image amplitude is the sum of the amplitude of all points in the spatial frequency

domain. And the total image amplitude was averaged among all screen shots belonging to a single experimental session. A cumulative image amplitude distribution profile is given in Figure 4.

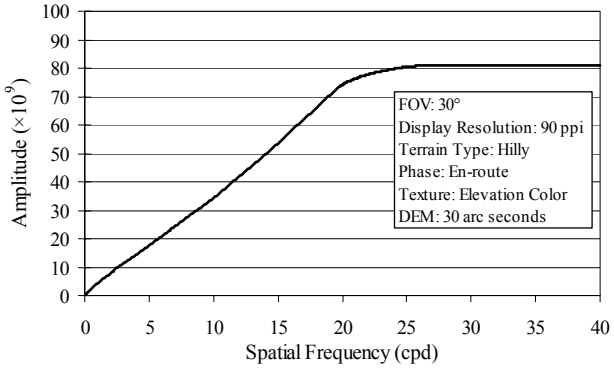


Figure 4. The Cumulative Spatial Amplitude Distribution for the Image in Figure 2 Filtered by a 90 ppi Low Pass Filter

Results

During the hilly flight path and terrain following phase, pilot control was mainly guided by the terrain information. The results are presented for XTE, total image amplitude, and inverse total image amplitude as a function of the display resolution as shown in Table 2. In Figure 5, averaged XTE and the logarithmic of inverse total image amplitude are plotted against the display resolution. It is obvious that average XTE and the logarithmic of inverse total image amplitude have a similar trend as functions of display resolution. When the display resolution was increased from 80 to 105 ppi, both average XTE and the logarithmic of inverse total image amplitude decreased. At 105 ppi, they reached the minimum level. However, when display resolution was increased from 105 to 120 ppi, average cross track error and inverse total image amplitude increased slightly as well.

In Figure 6, average XTE is plotted against the total image amplitude. A general trend was that the higher the total image amplitude, the smaller the average XTE. A logarithmic trendline was added in Figure 6. It seems that a logarithmic relation exists between XTE and the total image amplitude. Therefore, the total image amplitude was log transformed and plotted against average XTE in Figure 7. For the tested hilly and terrain-following scenarios listed in Table 1, the run sessions with 105 ppi had the highest average total image amplitude and the smallest average XTE. From these results, we conclude that higher levels of image amplitude will result in lower XTE (higher pilot performance). Higher resolution displays in general yield higher image amplitudes.

Table 2. XTE, Total Image Amplitude, and Inverse Total Image Amplitude Averaged by PPI

Display Resolution (ppi)	XTE (feet)	Average Total Image Amplitude ($\times 10^9$)	Average Inverse Total Image Amplitude ($\times 10^{-13}$)	Average Log_{10} Transform of Total Image Amplitude	Average Log_{10} Transform of Inverse Total Image Amplitude
80	832	46.063	217.092	10.663	-10.663
90	727	49.783	200.871	10.697	-10.697
105	501	188.167	53.144	11.275	-11.275
120	542	94.954	105.314	10.978	-10.978

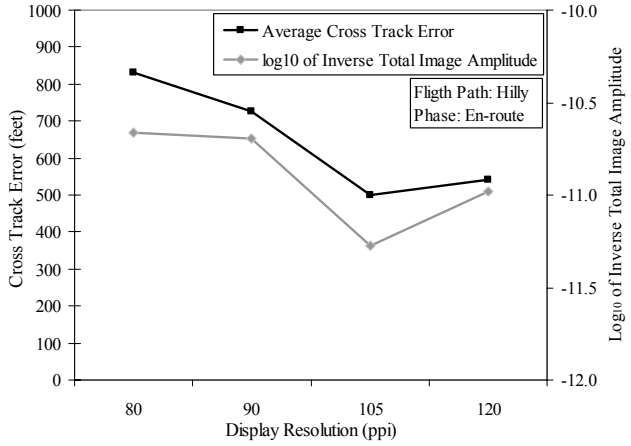


Figure 5. XTE and the Logarithmic of Inverse Total Image Amplitude Plotted as Functions of Display Resolution

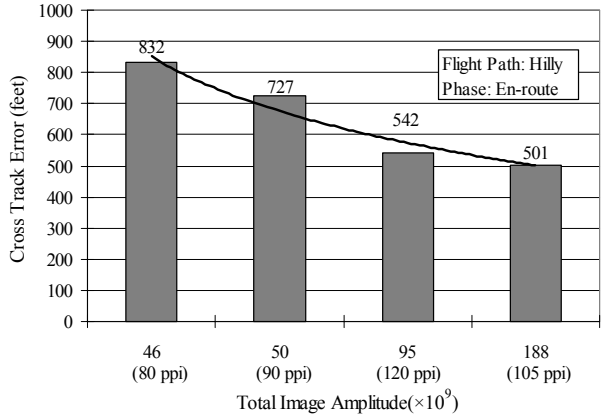


Figure 6. XTE Plotted against the Total Image Amplitude with a Logarithmic Trendline

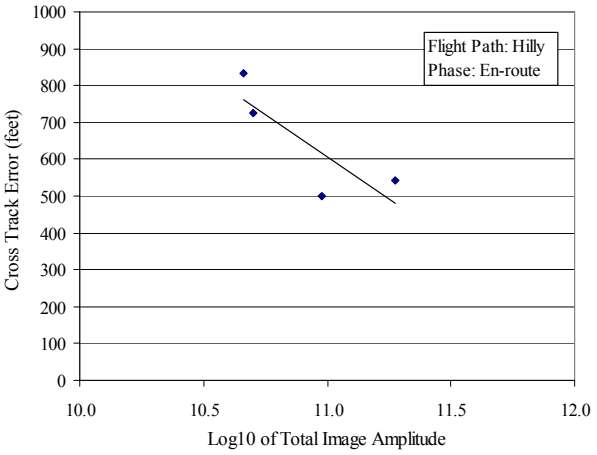


Figure 7. XTE plotted as a Function of the Logarithmic of the Total Image Amplitude with a Linear Trendline

XTE As A Function Of Field Of View (FOV)

We also investigated how the field of view (FOV) affected both pilot performance and image content. Refer to Table 3 for the results of this comparison. From Figure 8, Figure 9 and Figure 10, a strong pattern appears, illustrating that average XTE and the total image amplitude correlate well with SVS image FOV. The larger the FOV, the greater the amplitude and the smaller the XTE. From Figure 9, a logarithmic relation appears to exist between XTE and the total image amplitude

Table 3. XTE, Total Image Amplitude, and Inverse Total Image Amplitude Averaged by FOV

FOV (degrees)	XTE (feet)	Average Total Image Amplitude ($\times 10^9$)	Average Inverse Total Image Amplitude ($\times 10^{-13}$)	Average Log_{10} Transform of Total Image Amplitude	Average Log_{10} Transform of Inverse Total Image Amplitude
22	881	18.668	535.665	10.271	-10.271
30	678	77.178	129.570	10.887	-10.887
60	537	98.581	101.440	10.994	-10.994
90	506	184.541	54.189	11.266	-11.266

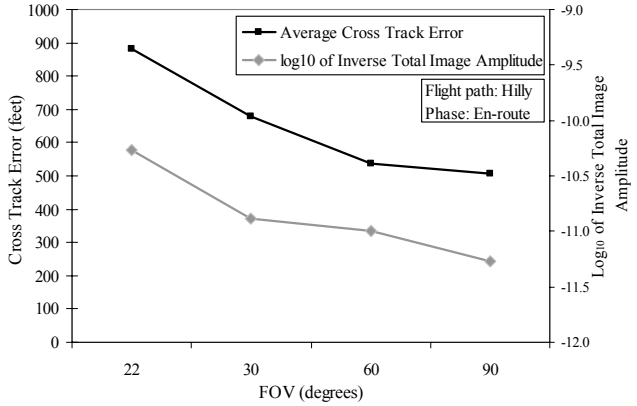


Figure 8. XTE and the Logarithmic of Inverse Total Image Amplitude Plotted as a Function of FOV

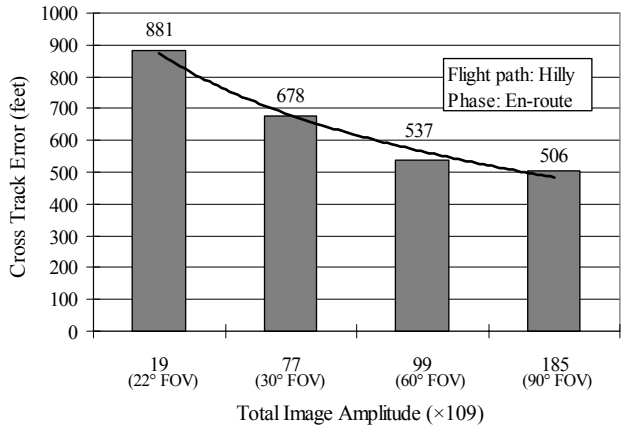


Figure 9. XTE Plotted against Total Image Amplitude with a Logarithmic Trendline

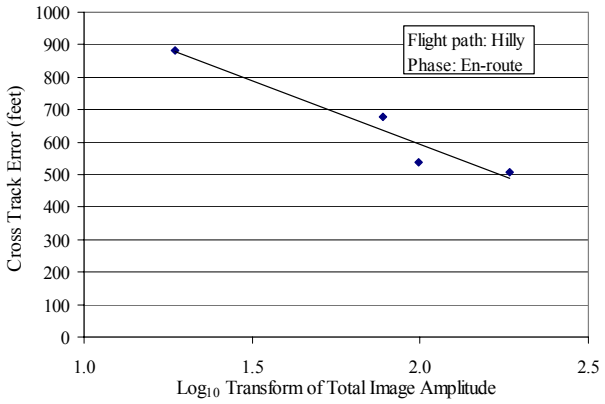


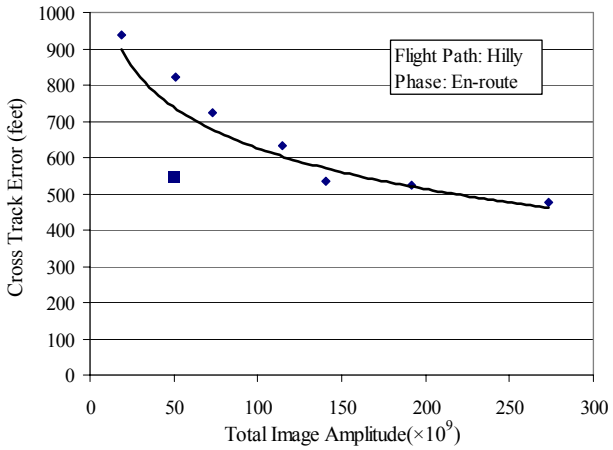
Figure 10.XTE Plotted against the Logarithmic of the Total Image Amplitude with a Linear Trendline

XTE As A Function Of Total Image Amplitude

Figure 11 shows a plot of the average XTE against total image amplitude for the hilly flight path and terrain-following phase scenarios. This figure indicates that a logarithmic relation might exist between average XTE and total image amplitude. A logarithmic transformation was performed on total image amplitude and plotted in Figure 12. We ran a regression model to obtain the trend line as shown in Equation (13).

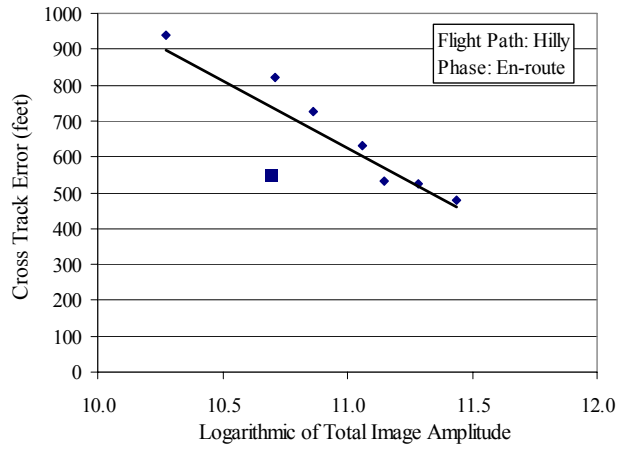
$$XTE = 5394.66 - 431.26 \log_{10}(total\_image\_amplitude) \quad (13)$$

where XTE is an estimated value of the RMS cross track error in hilly terrain and when flying with sole reference to an SVS image. The R^2 of Equation (13) was 96.48%.



Note: The square data point is highlighted as a susceptible outlier.

Figure 11.XTE Plotted as a Function of the Total Image Amplitude with a Logarithmic Trendline



Note: The square data point is highlighted as a susceptible outlier.

Figure 12.XTE Plotted as a Function of the Logarithmic of Total Image Amplitude with a Linear Trendline

Conclusions and Discussion

The flight scenarios researched in this paper were experimental sessions with hilly terrain and a terrain-following phase. Pilot performance was evaluated in terms of XTE as a function of display resolution (PPI) and field of view (FOV).

We found that the logarithmic of the total image amplitude was a statistically significant factor to explain average XTE when flying a terrain following task. Higher total image amplitude consistently yielded better pilot performance.

We feel that the methodology outlined in this paper may be useful to get a rough notion of the XTE when nothing other than a terrain image is available. By subjecting this terrain image to an FFT analysis, it is possible to predict XTE using Equation (13). We are fully aware that the regression model used in this methodology needs additional data points to be statistically more representative of a larger pilot population and a wider range of display characteristics. However, we feel that the method could be used in obtaining performance estimates of new prototype SVS displays that show terrain images long before any simulated or actual test flights are performed.

References

[1] Glaab, Louis J., Mohammad A. Takallu, 2002, Preliminary Effect of Synthetic Vision System Displays to Reduce Low-Visibility Loss of Control and Controlled Flight Into Terrain Accidents (2002-01-1550), Hampton, VA, NASA Langley Research Center and Lockheed Martin.

[2] Hemm, Robert, 2000, Benefit Estimates of Synthetic Vision Technology (NS002S1), Hampton, VA, NASA Langley Research Center.

[3] Schnell, Thomas, Sohel Merchant, 2001, Assessing Pilot Performance in Flightdecks Equipped with Synthetic Vision Information Systems. (Technical Report) Cedar Rapids, IA: Rockwell Collins.

[4] Schwartz, Steven H., 1999, Visual Perception, Second Edition, Appleton & Lange, Stamford, Conn.

[5] Pantle, Allan, 1973, Spatial Information Coding in the Human Visual System: Psychophysical Data, Proceedings of the National Aerospace Electronics Conference, pp. 304-308.

[6] Ginsburg, Arthur P., 1978, Visual Information Processing Based on Spatial Filters Constrained by Biological Data: Doctoral Dissertation, University of Cambridge, England.

[7] Ginsburg, Arthur P., 1980, Specifying Relevant Spatial Information for Image Evaluation and Display Design: An Explanation of How We See Certain Objects, Vol. 21, No. 3, Proceedings of the Society for Information Display (SID).

[8] Helander, Martin, 1988, Handbook of Human-Computer Interaction, Elsevier Scinece Publishing Company, Inc.

[9] Sinai, Michael J., J. Kevin Deford, Todd J. Purkiss, and Edward A. Essock, 2000, Relevant Spatial Frequency Information in the Texture Segmentation of Night-Vision Imagery, Vol. 4023, Proceedings of SPIE: Enhanced and Synthetic Vision.

[10] Bracewell, R.N., 2000, The Fourier Transform and Its Applications, Boston: McGraw Hill

[11] Phadke, M.S., 1989, Quality Engineering Using Robust Design, Englewood Cliffs, NJ: Prentice Hall.

[12] Gonzalez, Rafael C., Richard E.Woods, 2002, Digital Image Processing, Second Edition, Prentice Hall, Upper Saddle River, New Jersey.